



Династия

Фонд Дмитрия Зимина

# Школа молекулярной и теоретической биологии 2014

<http://dynastybioschool.wordpress.com>

HHMI  
HOWARD HUGHES MEDICAL INSTITUTE



## List of laboratories and projects / Список лабораторий и проектов:

Laboratory of Protein Biosynthesis, *head of laboratory Elena Alkalaeva* / Лаборатория биосинтеза белка, *руководитель Елена Алкалаева*.

Laboratory of Bacterial Genomics, *head of laboratory Mikhail Gelfand* / Лаборатория бактериальной геномики, *руководитель Михаил Гельфанд*.

Project 1. *ubiJ* or not *ubiJ* (*Zoya Chervontseva, Svetlana Petrova*) / Проект 1. *ubiJ* или не *ubiJ* (*Зоя Червонцева, Светлана Петрова*).

Project 2. Nitrogen? We'll fix you (*Jelena Chuklina, Svetlana Petrova*) / Проект 2. Азот? Ща зафиксируем (*Елена Чуклина, Светлана Петрова*).

Project 3. Leaderless... (*Jelena Chuklina*) / Проект 3. Без лидера (*Елена Чуклина*).

Project 4. Naked Ribosome (*Sofya Garushyants, Artur Zalevsky, Svetlana Petrova*) / Проект 4. Голая рибосома (*Софья Гарущянц, Артур Залевский, Светлана Петрова*).

Project 5. Sugar Lego (*Anna Kaznadzey, Zoya Chervontseva*) / Проект 5. Сахарное Лего (*Анна Казнадзей, Зоя Червонцева*).

Project 6. Dirty Daphnia (*Irena Artamonova*) / Проект 6. Грязная дафния (*Ирена Артамонова*).

Laboratory of Jose Garcia Perez, *head of laboratory Jose Luis Garcia Perez* / Лаборатория Хосе Гарсиа Переса, *руководитель Хосе Луис Гарсиа Перес*.

Laboratory of Bias Gene Conversion, *head of laboratory Guillaume Filion* / Лаборатория геной конверсии, *руководитель Гийом Фийон*.

Laboratory of Maxim Imaev, *head of laboratory Maxim Imaev* / Лаборатория Максима Имаева, *руководитель Максим Имаев*.

Laboratory of Vera Korneeva, *head of laboratory Vera Korneeva* / Лаборатория Веры Корнеевой, *руководитель Вера Корнеева*.

Laboratory of Vanya Kulakovskiy, *head of laboratory Vanya Kulakovskiy* / Лаборатория Вани Кулаковского, *руководитель Ваня Кулаковский*.

Laboratory Meer–Tutukina, *heads of laboratory Margo Meer and Maria Tutukina* / Лаборатория Меер–Тутукиной, *руководители Марго Меер и Мария Тутукина*.

Laboratory of Rational Drug Design, *heads of laboratory Peter Vlasov and Polina Shichkova* / Лаборатория Rational Drug Design, *руководители Пётр Власов и Полина Шичкова*.

## Laboratory of Protein Biosynthesis

*Head of laboratory: Elena Alkalaeva (IMB, Moscow, Russia)*



The goal of the project is to investigate experimentally the molecular mechanism of the protein biosynthesis process. One of the most unclear stages of protein synthesis in cells is translation termination. According to the newest results, many different proteins are the participants of the termination stage. For instance, besides two “canonical” and the well-known termination factors eRF1 and eRF3, in humans we have two others interesting proteins important for termination, Dbp5 and PABP. These proteins have a wide range of activities in the cells, the termination process is the additional one, but this important activity is not investigated well. During the project, we plan to obtain a set of mutant forms of these proteins and to test their activities in reconstructed system of protein synthesis (so-called “cell-free translation system”). We suppose to use the variety of modern methods of molecular biology, such as direct mutagenesis, cloning, recombinant proteins expression in bacteria and others.

## **Лаборатория биосинтеза белка**

*Руководитель: Елена Алкалаева (IMB, Moscow, Russia)*

Проект будет посвящен выяснению механизма заключительной стадии биосинтеза белков у человека (терминации трансляции). Дело в том, что это наименее изученная стадия синтеза белка в клетках. Как выяснилось в последние годы, в ней принимают участие намного больше белков, чем считалось раньше. Так нами недавно показано, что кроме факторов терминации eRF1 и eRF3 у человека имеется как минимум еще два белковых фактора Dbp5 и PABP, участвующих в этом процессе. Эти белки выполняют и другие функции в клетках, а участие в терминации трансляции является для них дополнительной активностью. Во время работы в летней школе, мы планируем получить мутантные формы этих белков и протестировать их активность в реконструированной в пробирке системе трансляции человека. Предполагается использовать широкий спектр методов молекулярной биологии: сайт-направленный мутагенез, клонирование, экспрессия рекомбинантных белков в клетках бактерий и их очистка и т.д.

## Laboratory of Bacterial Genomics

*Head of laboratory: Mikhail Gelfand (MSU, Moscow, Russia)*



**Project 1. *ubij* or not *ubij*** (*Zoya Chervontseva, Svetlana Petrova*)

When molecular biologists reach an impasse and there is a contradiction between the results of different labs, bioinformatics can help. In the region 4 019 624 – 4 020 229 of the *E.coli* genome some researchers see a protein-coding gene and while others insist that it encodes a small non-coding (likely, regulatory) RNA. This gene is assumed to be involved in the ubiquinone biosynthesis, but no functional details are known. We will use the comparative genomic approach to investigate this case. This is a project for 1-2 students.

## Лаборатория бактериальной геномики

Руководитель: Михаил Гельфанд (МГУ, Москва, Россия)

### Проект 1. *ubiJ* или не *ubiJ* (Зоя Червонцева, Светлана Петрова)

Когда молекулярные биологи заходят в тупик, и результаты разных лабораторий противоречат друг другу, может помочь биоинформатика. В районе 4 019 624 – 4 020 229 генома кишечной палочки одни исследователи видят белок-кодирующий ген *ubiJ*, а другие доказывают, что здесь закодирована некодирующая (видимо, регуляторная) РНК. Считается, что функция гена связана с синтезом убихинона в аэробных условиях, но как именно – неизвестно. В этом проекте мы проведём сравнительно-геномный анализ этого участка и постараемся понять, что же происходит на самом деле. Это проект для 1-2 участников.

## **Project 2. Nitrogen? We'll fix you** (*Jelena Chuklina, Svetlana Petrova*)

**Reconstruction of the nitrogen fixation regulon in soybean symbiotic bacterium *Bradyrhizobium japonicum* by integration of comparative genomic, dRNA-seq and microarray data.** Nitrogen is the most abundant gas of atmosphere and nitrogen atoms are an essential part of many biologically important macromolecules, including proteins, DNA and RNA. However, only few bacteria are able to convert atmospheric nitrogen to a biologically consumable form. Some of these bacteria perform nitrogen fixation in symbiosis with plants, which in return provide the bacteria with nutrients. We aim to understand how bacteria regulate this process and what components (proteins, small non-coding RNA) are regulated. We have accumulated high quality data for soy symbiont *Bradyrhizobium japonicum* from various sources: motifs that transcription factor recognize in DNA (which point to genes possibly regulated by these factors), microarray data (telling which regions of the genome are active during nitrogen fixation) and dRNA-seq data (exactly mapping transcription starts). The goal is to integrate these data and find new regulated genes and snRNAs. Project can be carried out in parallel by 2-3 students.



## Проект 2. Азот? Ща зафиксируем (Елена Чуклина, Светлана Петрова)

**Интеграция данных сравнительной геномики, микрочипов и dRNA-seq для реконструкции транскрипционной регуляции азотфиксации симбионта сои *Bradyrhizobium japonicum*.** Азот составляет около 70% атмосферы, а также входит в состав многих биологических макромолекул, таких как белки и нуклеиновые кислоты (ДНК и РНК). Однако, молекулярный азот - химически очень устойчивое соединение, и превращать его в биологически доступную форму способны только некоторые бактерии. Некоторые из этих бактерий фиксируют азот в симбиозе с растениями (чаще всего бобовыми), а растения, в свою очередь, делятся с бактериями питательными веществами. Мы хотим понять, как бактерии регулируют этот сложный и энергетически дорогой процесс, и, что важно, какие компоненты (белки, малые некодирующие РНК) регулируются совместно с генами азотфиксации. К настоящему моменту для одной такой бактерии *Bradyrhizobium japonicum*, фиксирующей азот в симбиозе с соей, накоплено большое количество данных: распознающие правила для сайтов связывания транскрипционных факторов (которые указывают, какие участки генома регулируются этими факторами, и, тем самым, какие гены они регулируют), микрочипы (которые показывают, какие участки генома активны в условиях азотфиксации), и данные dRNA-seq (показывают старты транскрипции в условиях азотфиксации, а регулируемые участки почти всегда расположены около стартов транскрипции). В ходе проекта мы обобщим эти данные, чтобы найти ранее неизвестные гены, которые включаются при фиксации азота. Проект рассчитан на параллельное выполнение 2-3 участниками школы.

### **Project 3. Leaderless... (Jelena Chuklina)**

**Leaderless mRNA of *B.japonicum* inferred from dRNA-seq.** The dRNA-seq data of bacterium *Bradyrhizobium japonicum* identifies transcription start sites at a single-nucleotide resolution. The method is quite new (the first article was published in 2010) and it is not clear how much information can be found in these data. One interesting question is how common are leaderless mRNAs in bacteria. Normally mRNA has a leader sequence, i.e. 20-40 nucleotides between the transcription and translation start sites, so that the ribosome has some space to bind. However, in leaderless transcripts, the transcription and translation start at the same point. The goal of the project is to assess, how many of mRNAs are indeed leaderless and how many are likely mRNA degradation products. After leaderless mRNAs are identified, we want to test, whether we find a special promoter motif upstream (we have good reasons to believe there is one). The project requires 1 student. Interest to learn some Python programming could benefit the project, but is not essential.

### Проект 3. Без лидера (Елена Чуклина)

**Изучение безлидерных мРНК, найденных в dRNA-seq *Bradyrhizobium japonicum*.** Данные dRNA-seq для бактерии *Bradyrhizobium japonicum* позволяют определять старты транскрипции с точностью до нуклеотида. Метод dRNA-seq достаточно новый (опубликован в 2010 году), и пока неясно, насколько много информации можно извлечь из экспериментальных данных такого типа. В частности, эти данные позволяют оценить, насколько распространены безлидерные мРНК. Обычно у каждой мРНК есть лидер, или последовательность длиной 20-40 нуклеотидов между стартами транскрипции и трансляции, чтобы рибосома могла связаться с мРНК. В безлидерных же мРНК старты транскрипции и трансляции совпадают. В этом проекте мы хотим узнать, сколько мРНК действительно начинаются со старта трансляции, а сколько являются продуктами распада обычной мРНК с лидерной последовательностью. Мы составим список безлидерных мРНК и проверим, действительно ли эти мРНК транскрибируются с промотора особого типа (мы предполагаем, что это именно так). Интерес к изучению или владение языком программирования Python ускорит и упростит выполнение проекта, но не является обязательным условием. Проект рассчитан на одного человека.

**Project 4. Naked Ribosome** (*Sofya Garushyants, Artur Zalevsky, Svetlana Petrova*)

Endosymbiotic bacteria help mealybugs and aphids to survive on the plant sap diet lacking essential amino acids. While endosymbiotic bacteria in insects have belong to different taxonomic groups, all of them are unable to survive outside the host and show signs of drastic genome degradation, such as accumulation of pseudogenes, gene loss, and deletion of the large genome regions. At least 10 Gammaproteobacteria genomes on different stages of degradation have sequenced from various insect hosts. The goal of this project is to study the minimal translational apparatus in these reduced-genome bacteria and to characterize changes in the protein composition and structure of their ribosomes. This is a project for 1-3 students.

#### **Проект 4. Голая рибосома** (*Софья Гарущяни, Артур Залевский, Светлана Петрова*)

**Определение минимального состава рибосомы у бактерий-эндосимбионтов насекомых.** Бактерии-эндосимбионты позволяют сосущим насекомым, таким как тли и червецы, питающимся только растительным соком, получать необходимое количество незаменимых аминокислот. Такие бактерии относятся к различным таксономическим группам, но все они неспособны выжить вне организма-хозяина. Плотная ассоциация этих бактерий с насекомыми приводит к деградации их геномов, в частности, к накоплению псевдогенов, уменьшению общего количества генов и потере целых фрагментов генома. В настоящий момент известны последовательности как минимум 10 геномов эндосимбионтов, относящихся к гамма-протеобактериям и находящихся на разных стадиях деградации. Целью этого проекта является изучение особенностей аппарата трансляции у таких бактерий, в первую очередь изменения белкового состава и структуры рибосомы. Это проект для 1-3 участников.

**Project 5. Sugar Lego** (*Anna Kaznadzey, Zoya Chervontseva*)

**Sugar metabolism genes: evolution, cooperation and Lego constructor.** Bacteria are able to metabolize many diverse sugars and grow on highly variable media. This ability depends on specific genes. These genes are responsible for transporting sugar into the cell, splitting long carbohydrates into shorter ones, catalyzing various chemical reactions, etc. We want to understand how these genes are arranged in the bacterial chromosome, which genes are the most popular in nature, which of them tend to be neighbors, i.e. form “gene cassettes”, how often genes from such cassettes interchange among different bacteria like Lego constructor pieces and why. We will specifically look into several cases regarding the most abundant metabolic pathways in the bacterial world. This project will not only help to explore carbohydrate metabolism, but get general idea on how diverse bacteria are, how bacterial genetics and horizontal transfer works and much more. The project requires 2-3 students.

## **Проект 5. Сахарное Лего** (*Анна Казнадзей, Зоя Червонцева*)

### **Гены сахарного метаболизма: эволюция, кооперация и конструктор Лего.**

Бактерии способны усваивать самые разнообразные сахара и расти, соответственно, на самых разных средах. Для этого у них имеются гены, отвечающие за разные этапы этого усваивания: транспорт веществ в клетку, расщепление крупных углеводов на более мелкие, всевозможные химические превращения и т.п. Мы хотим посмотреть, как эти гены расположены на бактериальной ДНК, какие из них наиболее востребованы в природе, какие из них предпочитают быть соседями (составляя "касеты" генов), в каких случаях гены таких кассет тасуются между разными бактериями, подобно строительным элементам Лего, и зачем. Отдельно мы разберем несколько конкретных примеров, касающихся самых распространенных в бактериальном мире метаболических путей. Эта задача позволит не только разобраться в углеводном метаболизме, но и получить общие представления о разнообразии и генетике бактерий, горизонтальном транспорте генов и многом другом. Проект рассчитан на 2-3 участников.

## **Project 6. Dirty Daphnia** (*Irena Artamonova*)

**Bacterial “contamination” of recently sequenced metazoan genomes.** Some genes predicted for newly sequenced metazoan genomes usually seem to be much more similar to bacterial genes than to genes of evolutionarily closer, eukaryotic organisms. Are they results of horizontal gene transfer or simple contamination of experimental samples? Or maybe they belong to bacterial symbionts of target organisms? We will try to answer such questions for bacterial-like genes of the Ctenophore, the largest animal that uses cilia to swim, and Daphnia, a small planktonic crustacean, by analyzing DNA features of genome regions where these genes reside. This is a project for 1-2 students.



## **Проект 6. Грязная дафния** (*Ирена Артамонова*)

**Бактериальное “загрязнение” в процессе секвенирования геномов животных.** В процессе описания генов, предсказанных для вновь прочитанных геномов животных, часто выясняется, что небольшая часть этого множества значительно больше похожа на гены бактерий, чем на гены, встречающиеся в эволюционно более близких, эукариотических, организмах. Как такие гены попали в геном животного? Может быть, это артефакт эксперимента по подготовке генома к чтению? Оказывается, ответить на этот вопрос можно, если проанализировать участки ДНК, содержащие такие гены – особенности, характерные для протяженной нуклеотидной последовательности, позволяют отличить бактериальную ДНК от эукариотической, и даже описать эволюционную историю генов. В этом проекте мы исследуем бактериальные гены, найденные при прочтении геномов гребневика, самого большого морского животного среди передвигающихся при помощи ресничек, и дафнии, представителя планктонных ракообразных. Мы попытаемся отличить гены, встроившиеся в геном в результате горизонтального переноса, от генов бактериальных симбионтов или загрязнений экспериментальной пробы. Это задача для 1-2 участников.

## Laboratory of Jose Garcia Perez

*Head of laboratory: Jose Luis Garcia Perez (GENYO, Granada, Spain)*



Surprisingly, only 5% of the human genome is made of genes. Perhaps even more of a surprise is that almost half of our genome is made of repetitive DNA sequences. Among them, some pieces of repeated DNA have the intrinsic property of mobilizing within the genome, leading to its expansion in size. As their movement is random in the genome, new insertions can lead to a human genetic disorder due to the disruption of a gene. In the project, we will conduct assays in cultured cells to test how often and how do repeated DNA move in our genome.

## **Лаборатория Хосе Гарсиа Переса**

*Руководитель: Хосе Луис Гарсиа Перес (GENYO, Granada, Spain)*

Удивительно, но всего лишь 5% генома человека состоит действительно из генов. Пожалуй, еще более удивительно, что почти половина нашего генома состоит из повторяющихся участков ДНК. Среди таких повторяющихся участков ДНК некоторые имеют индивидуальную способность перемещаться внутри генома. Это приводит к увеличению размера генома. Поскольку такие перемещения в геноме совершенно случайны, вставки новых последовательностей могут приводить к генетическим заболеваниям (перестройкам), которые связаны с нарушением нуклеотидного состава гена. В предлагаемом проекте мы будем проводить анализ культивируемых клеток, чтобы проверить, как часто происходят перемещения повторяющихся участков ДНК в нашем геноме.

## Laboratory of bias gene conversion

*Head of Laboratory: Guillaume Filion (CRG, Barcelona, Spain)*



**The purpose of this class is to test whether chromosome breaks are repaired similarly when they happen at different places.** In the nucleus, chromosome breaks can be repaired in two major ways. In the first, called Non Homologous End Joining, the DNA ends are simply put back together. In the second, called Homologous Recombination, the repair process is more complex and involves exchange of DNA between chromosomes. How does the cell choose which repair pathway to use? More specifically, does the *location* of the break matter for the choice and for repair events? The participants will answer this question with standard molecular biology experiments (PCR, restriction enzyme digestion, agarose gel electrophoresis). Part of the course will also involve cell culture and fluorescence microscopy.

## **Лаборатория генной конверсии**

*Руководитель: Гийом Фийон (CRG, Barcelona, Spain)*

**В рамках проекта участникам предстоит проверить, одинаков ли механизм репарации ДНК-разрывов, локализованных на разных участках хромосом.**

Восстановление нативной структуры ДНК после хромосомных разрывов в клеточном ядре может быть реализовано посредством двух основных механизмов. Первый называется «соединение негомологичных ДНК-концов» и заключается в простом «сшивании» двух фрагментов поврежденной ДНК. Второй – гомологичная рекомбинация – более сложный механизм, так как он предполагает обмен ДНК между хромосомами.

Как клетка “решает” какой из путей ДНК-репарации ей выбрать, и может ли положение разрыва на хромосоме повлиять на это “решение”? Ответить на эти вопросы участники проекта смогут с помощью стандартных методов молекулярной биологии (ПЦР, рестрикция ДНК, электрофорез в агарозном геле). Отдельная часть проекта также будет включать основы культивирования клеток и флуоресцентной микроскопии.

## Laboratory of Maxim Imakaev

*Head of laboratory: Maxim Imakaev (MIT, Cambridge, USA)*



Our lab will use polymer simulations and genomic analyses to investigate 3D organization of chromosomes. Chromosome organization plays critical roles for a range of cellular processes, including: gene regulation, DNA replication, and chromosomal alterations in cancer cells. In 2009, the Hi-C method was introduced to study chromosomes on a genome-wide scale and has since provided many important insights. However, Hi-C provides data about chromosomal contacts in a population of cells, instead of spatial positions of chromosomal regions in each cell, making it challenging to interpret. Over the last four years, we have been developing the following approach to analyze this data: first, we build polymer models of chromosomes based on known or hypothesized principles of chromosome organization; we then compare these models to Hi-C and other (e.g. microscopy) experimental data; finally, we select those which best agree with the experimental data. At the school, we will take a similar approach to study questions regarding chromosome organization and dynamics, including: 1) whether attraction to the nuclear lamina may lead to formation of chromosomal domains in the Hi-C data. 2) whether contact probability measured by Hi-C can be interpreted as the time it takes for the two genomic elements to find each other. 3) whether binding of microscopy probes to DNA may change the way DNA diffuses, thus introducing biases to results of the microscopy measurements. You will choose the project when you join the lab; we also have several more projects in mind, from polymer simulations to data analyses. The projects mentioned above do not require programming experience; however, if you do know Python programming, and want to learn advanced data analyses techniques, we can develop a project to best suit you.

## **Лаборатория Максима Имакаева**

*Руководитель: Максим Имакаев (MIT, Cambridge, USA)*

Наша лаборатория будет заниматься полимерными симуляциями и анализом геномных данных в области трехмерной организации хромосом. Организация хромосом играет важную роль во многих процессах внутри клетки, таких как регуляция генов, репликация ДНК, и хромосомные перестройки в раковых клетках. В 2009 году был предложен новый метод для изучения хромосом - "Хай-Си" (Hi-C). Этот метод позволяет изучать структуру всего генома в одном эксперименте; за последние 5 лет с помощью этого метода было сделано много интересных открытий. Однако, анализировать данные Хай-Си очень сложно, поскольку этот метод измеряет структуру контактов в среднем по целой популяции клеток, тогда как мы хотели бы знать позиции разных районов генома в каждой клетке. В последние 4 года, наша лаборатория разрабатывала особый подход к анализу данных Хай-Си: сначала мы строим полимерные модели хромосом, основываясь на известных и гипотетических принципах устройства хромосом. Затем мы симулируем Хай-Си данные на основании полученных полимерных моделей, и отбираем модели, которые согласуются с экспериментальными данными. На нашей школе мы планируем использовать этот метод для изучения следующих вопросов об организации хромосом: 1) может ли притяжение хромосом к оболочке ядра привести к образованию доменов хроматина, наподобии тех, что видны в Хай-Си данных. 2) можно ли интерпретировать данные Хай-Си, как время, необходимое для того чтобы два геномных элемента встретились в ядре 3) могут ли флюорофоры, используемые в микроскопии живых клеток, повлиять на динамику ДНК, и тем самым внести ошибку в экспериментальные измерения. Вы можете выбрать конкретный проект после того, как вы присоединитесь к нашей лаборатории; кроме этих трех проектов у нас есть еще проекты, от полимерных симуляций до проектов по анализу данных. Все перечисленные проекты не требуют умения программировать; однако, если Вы умеете программировать на Python и хотите научиться современным методам анализа данных, то мы найдем проект который подходит именно вам.

## **Laboratory of Vera Korneeva**

*Head of laboratory: Vera Korneeva (HSC, Moscow, Russia)*



As it is in our real life, we decided to make a "wet-dry" lab, where experiment and theory is combined. Our main goal is to show that Biological sciences use quantitative description, mathematical models and precise laws. This year we will continue with our project on investigation and characterization of serine protease inhibitors. "Wet" part of the project in our laboratory will include isolation and purification of serine proteases (trypsin) and their inhibitors from natural sources, as well as the study of their interaction by biophysical methods. During the "dry" a mathematical model of the interaction between the isolated enzymes and inhibitors will be developed in order to determine the type of this interaction.



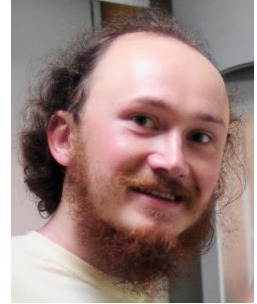
## **Лаборатория Веры Корнеевой**

*Руководитель: Вера Корнеева (ГНЦ, Москва, Россия)*

В соответствии со стилем нашей собственной научной деятельности мы решили сделать «мокро-сухую» лабораторию, где будет сочетаться эксперимент и теория. Наша главная цель заключается в том, чтобы показать, как в любой части биологической науки могут быть количественные описания, математические модели и точные законы. В этом году мы продолжим работу над поиском и характеристикой ингибиторов сериновых протеаз. «Мокрая» часть работы в нашей лаборатории будет включать выделение и очистку сериновых протеаз (трипсина) и их ингибиторов из природного сырья, а так же изучение их взаимодействия методами биофизики. Во время «сухой» части нашей работы будет построена математическая модель взаимодействия выделенных ферментов и ингибиторов, что позволит определить тип взаимодействия.

## Laboratory of Vanya Kulakovskiy

*Head of laboratory: Vanya Kulakovskiy (IOGen, Moscow, Russia)*



We will focus on several semi-independent projects in the field of computational genomics and transcriptomics of eukaryotes.

**Project 1.** Sequence read quality check, trimming, mapping etc. are all important step to biological interpretation of genomic and transcriptomic data. Usually the base analysis is performed using existing command-line software through remote access to high-performance computational resources (but not at your own laptop). This project will mainly focus on human and plant transcriptomes from the very beginning "raw reads" processing.

Required skills: advanced user-level computer skills. Experience with Linux environment and/or basic programming skills will be helpful.

**Project 2.** Detailed analysis of poorly annotated and unannotated genes, activated in yeast Mn<sup>2+</sup> stress response. Preliminary transcriptomic data for yeast manganese stress (as a model for heavy metal toxicity) revealed dozens of activated genes with unannotated or unknown gene function. Using existing bioinformatics software and databases we shall try digging potential function of these genes.

Required skills: knowledge of basic molecular biology. Erudition in molecular biology will be helpful. Programming skills are not necessary.

**Project 3.** Comparative analysis of transcription start sites of mTOR targets in human and mouse genomes.

This projects is focused on sequence analysis of transcription initiation regions of genes, translationally regulated by mTOR. This projects follows the results of previous schools (2012-2013).

Previously we found that several broad transcription start sites of mTOR-targets contained both pyrimidine-rich and pyrimidine-depleted regions, thus potentially producing mRNA isoforms specifically regulated and not regulated by mTOR. This year we are going to compare "transcriptional activity profiles" of mTOR targets in human and mouse genomes.

Required skills: basic programming skills are necessary. Knowledge of any scripting language (Ruby, Python, Perl) will be helpful.

## Лаборатория Вани Кулаковского

*Руководитель Ваня Кулаковский (ИОГен, Москва, Россия)*

Работа в лаборатории будет построена вокруг трех независимых проектов в области вычислительной геномики и транскриптомики эукариот.

**Проект 1.** Компьютерный анализ данных современного высокопроизводительного секвенирования для сборки геномов, транскриптомов и исследования дифференциальной экспрессии генов высших эукариот. Данные, полученные с секвенаторов нового поколения, представляют собой миллионы коротких нуклеотидных последовательностей, "чтений". Компьютерная обработка (контроль качества, обрезка, картирование на геном и т.д.) чтений является важным шагом на пути к биологической интерпретации геномных и транскриптомных данных и, как правило, выполняется существующими вычислительными программами в командной строке. Достаточно часто используются не локальные вычислительные ресурсы (твой домашний ноутбук), а удаленный доступ к мощным вычислительным машинам. В рамках проекта мы будем заниматься анализом растительных и человеческих транскриптомов, в том числе "с нуля", т.е. начиная от базовой обработки "сырых чтений".

Необходимые навыки: свободное владение компьютером на уровне пользователя. Плюсом будет знакомство с операционной системой Linux и навыки программирования.

**Проект 2.** Тщательный анализ неаннотированных и плохо аннотированных генов, активирующихся у дрожжей при марганцевом стрессе. Предварительный взгляд на транскриптомные данные дрожжей, выращенных на среде с большим избытком марганца (модель отравления тяжелыми металлами), показал несколько десятков активировавшихся генов, функция которых не известна или по каким-то причинам не аннотирована в основной базе данных дрожжевых генов. С помощью существующих программ и баз данных, созданных для биологов и доступных через сеть интернет, мы будем пытаться разобраться с возможной функцией этих генов.

Необходимые навыки: знание основ молекулярной биологии. Широкая эрудиция в молекулярной биологии будет плюсом. Навыки программирования не потребуются.

**Проект 3.** Сравнительный анализ сайтов инициации транскрипции у генов-мишеней сигнального пути mTOR у человека и мыши. Проект посвящен анализу районов инициации транскрипции генов, специфически регулируемых белком mTOR на уровне трансляции. Проект является продолжением работы, сделанной в рамках предыдущих школ.

Специфический ответ на mTOR на уровне трансляции обеспечивается наличием пиримидин-богатого тракта в районе старта транскрипции. Ранее мы обнаружили, что в геноме человека широкие старты транскрипции многих mTOR-мишеней включают как пиримидин-богатые, так и пиримидин-бедные районы, то есть потенциально продуцируют изоформы мРНК, отвечающие и не отвечающие на mTOR. В этом году мы планируем сравнить "профиль активности" промоторов mTOR-мишеней у человека и мыши.

Необходимые навыки: базовый уровень программирования на каком-либо языке. Программирование на скриптовом языке (Ruby, Python, Perl) будет плюсом.

## Laboratory Meer-Tutukina

Heads of laboratory: Margo Meer (CRG, Barcelona, Spain) and Maria Tutukina (ICB, Pushchino, Russia)



Duplication of DNA is a facile method of evolution providing great opportunities for new gene functions to appear – it is easier to modify a functional gene than to create *de novo* by random mutations. This makes gene duplication to be one of the most exiting topics in molecular evolution.

To study evolutionary events which happened a long time ago, such as old duplications, we can use information on genomes of different species. But if we want to answer the questions related to recent events we need to have data on polymorphisms among the genomes of the same species.



The goal of our project is to obtain and analyze data on polymorphisms in the genes which were recently duplicated. To do so we will sequence and compare these genes from 12 individuals of *Schizophyllum commune*.

To implement this project we will use both types of methods: experimental (DNA extraction, purification on Qiagen columns and magnetic beads, analysis with agarose gel and Qubit, PCR, qPCR, library preparation for sequencing on MiSeq) and computational (work with databases, sequencing results analysis).

## Лаборатория Меер–Тутукиной

Руководители: Марго Меер (CRG, Barcelona, Spain) и Маша Тутукина (ИБК, Пущино, Россия)

Дупликация генов является одним из основных механизмов эволюции, создавая условия для возникновения новых функций – легче модифицировать уже работающий ген, чем случайными мутациями создавать его на пустом месте. Это делает дупликации интересным объектом исследований. В настоящее время существует большое количество информации, позволяющей изучать случаи давних дупликаций – тех, которые произошли в общих предках существующих ныне видов. Для этого достаточно сравнивать дупликации между разными видами. Однако для изучения недавних эволюционных событий такой анализ подходит плохо. В этом случае необходима не обобщённая информация о виде в целом, а сведения о вариациях в геномах отдельных индивидуумов - полиморфизмах. Целью нашего проекта будет получение и анализ данных о полиморфизмах в недавно дуплицированных генах. Для этого мы отсекуем и сравним такие гены у 12 представителей вида *Schizophyllum commune*. В проекте будут использованы как экспериментальные методы (выделение ДНК, PCR, qPCR, анализ ДНК в агарозном геле и с помощью Qubit, очистка ДНК с помощью колонок Qiagen и магнитных шариков, подготовка образцов к секвенированию методом MiSeq), так и биоинформатические (использование баз данных, анализ результатов секвенирования).

## Laboratory of Rational Drug Design

*Head of laboratory: Peter Vlasov (CRG, Barcelona, Spain) and Polina Shichkova (MIPT, Moscow, Russia).*



The main goal of the project is to present to the participants the concept of the modern rational drug design and hands-on experience of the corresponding methods and tools on the samples of search (prediction) of perspective ligands for actual, therapeutically perspective protein targets.

The educational process will include the theoretical lectures: general principles of protein structural organization and diversity of protein structures and functions; prediction of protein properties and folds by homology; physics and modeling of protein-ligand interactions; some “hot” topics of molecular biology closely connected with drug-design: RNA interference, molecular mechanisms in immune system.

The skills students will get in the project: experience in a modern professional software packages for modeling of protein structures and protein-ligand interactions; using various



Internet resources and databases of modern biology – genes, proteins, small molecules, drugs and others.

The presumable results: the project team will predict specific low-molecular compounds as the new potential ligands/regulators of several therapeutically perspective target proteins.

## **Лаборатория Rational Drug Design**

*Руководители: Пётр Власов (CRG, Barcelona, Spain) и Полина Шичкова (МФТИ, Москва, Россия).*

Главная цель проекта – дать «студентам» (участникам Школы) представление о современных методах разработки лекарственных препаратов на примере поиска (предсказания) перспективных лигандов для актуальных, терапевтически важных белков-мишеней.

Образовательная составляющая проекта включает теоретические лекции: общие принципы устройства и функционирования белков; предсказание свойств и структур белков; физика и моделирование взаимодействий белков с лигандами; некоторые «горячие темы» молекулярной биологии, близкие к drug-design (такие как RNA-интерференция, механизмы работы иммунной системы и пр.).

Научная составляющая включает изучение и использование современных методов моделирования белковых структур и белок-лигандных взаимодействий, с применением актуальных биологических ресурсов и баз данных – по генам, белкам, низкомолекулярным соединениям, лекарственным препаратам и пр.

Предполагаемый результат: предсказание новых веществ-лигандов, потенциально регулирующих активность терапевтически важных белков-мишеней.

## **Directors / Дирекция**

Anna Piotrovskaya – Executive Director of Dynasty Foundation / Анна Пиотровская  
Исполнительный директор Фонда Дмитрия Зимина «Династия»

Dmitry Kokorin – Director of Bioschool / Дмитрий Кокорин Директор школы

Fyodor Kondrashov – Scientific Director of Bioschool / Фёдор Кондрашов Научный  
директор школы

Victor Toldov – Project Manager of Dynasty Foundation / Виктор Толдов Менеджер  
проектов Фонда Дмитрия Зимина «Династия»